

MIIC-2 Kickoff

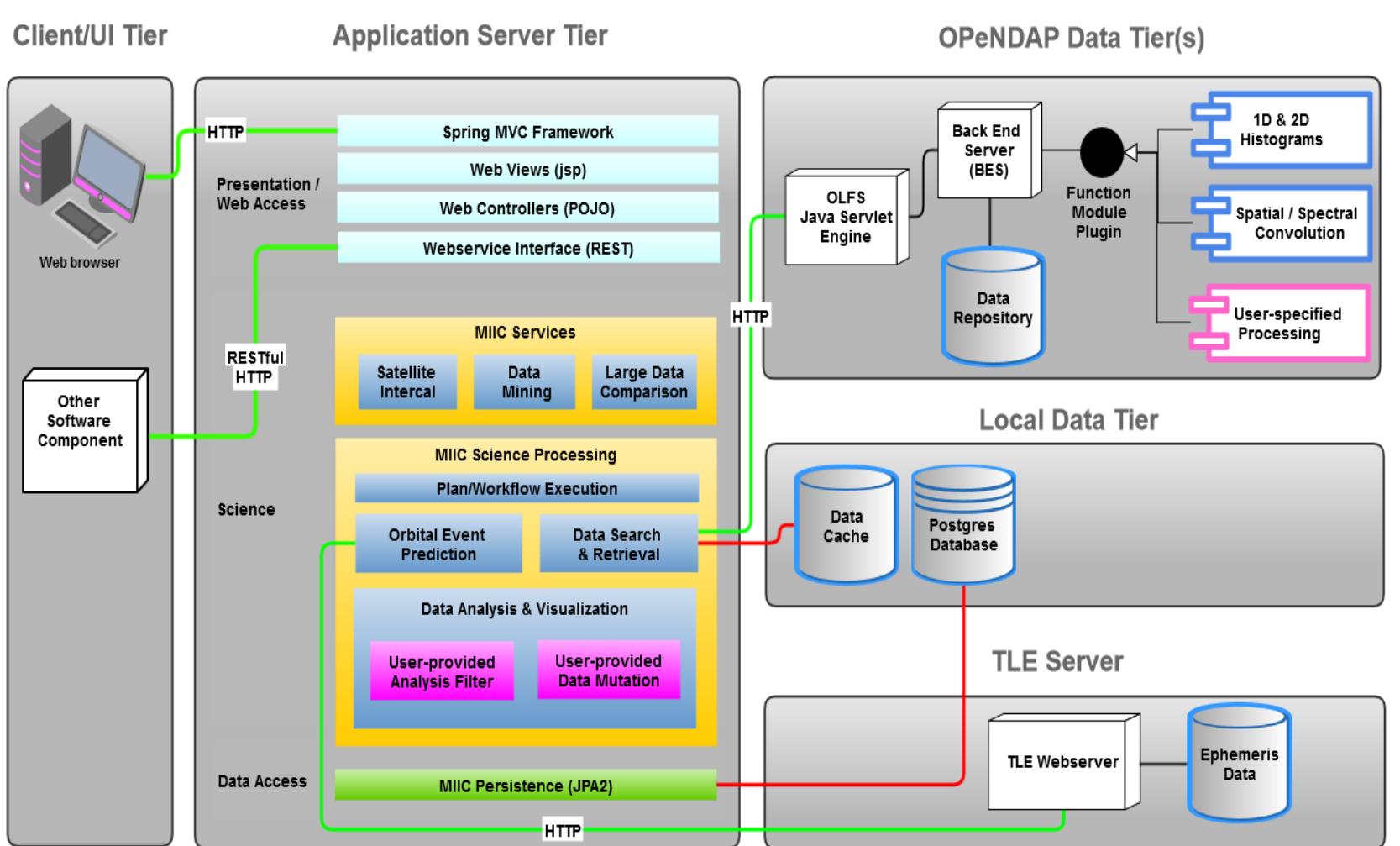
- Objectives of Meeting
- Agenda
- Background on Intercomparison
- MIIC-1 Status
- MIIC-2 Requirements

May 27, 2014

Objectives of Meeting

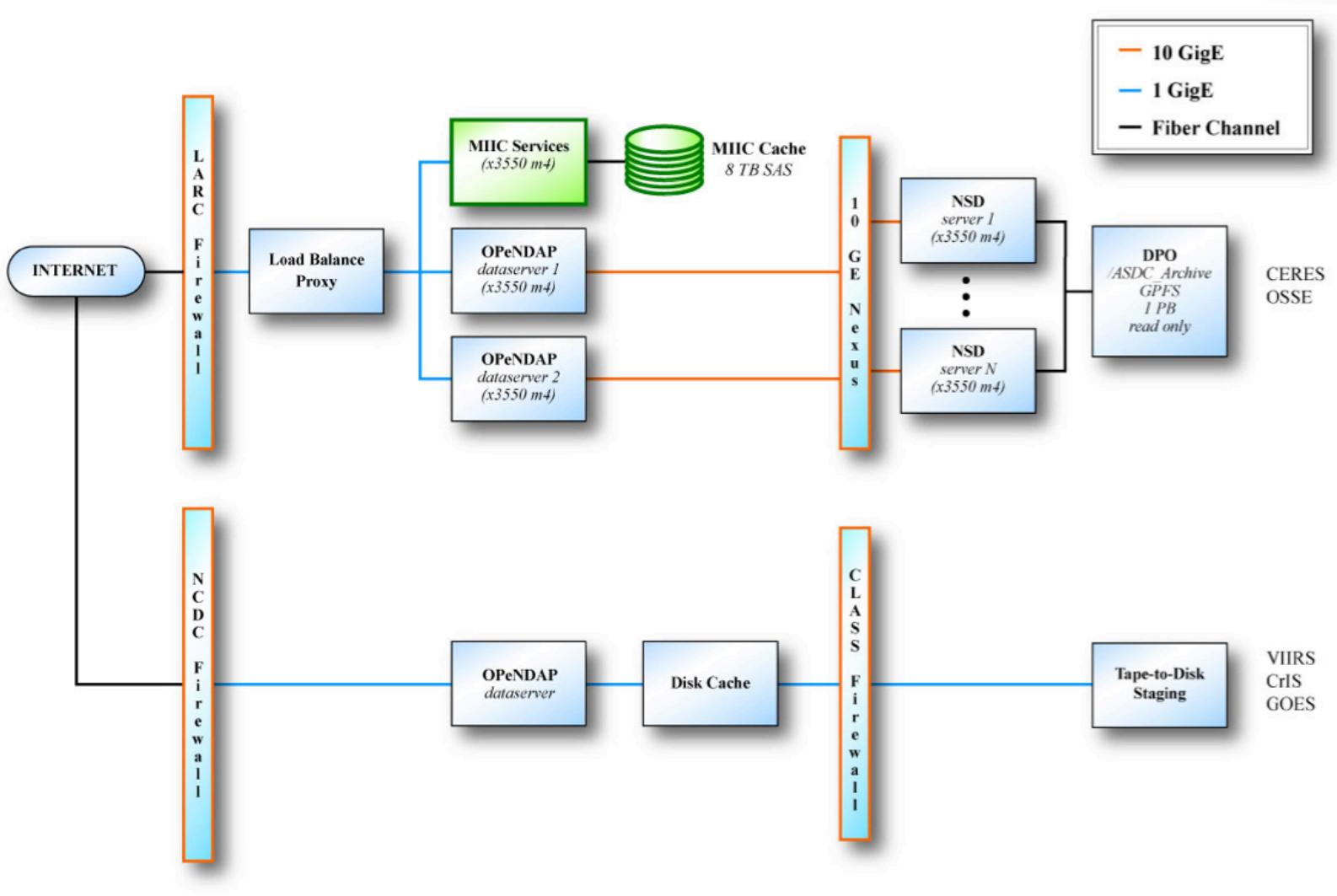
- Team building
 - Deal with geographic and agency differences
 - Respect data center boundaries
 - Leverage strengths of team members and facilities
 - Understand specifically how MIIC can benefit each center
 - Understand roles and responsibilities of team members
 - Use Collaborative Management (CM) tools
- Meeting artifacts
 - List key milestones
 - Identify issues/challenges
 - Scrub schedule presented to Curt Tilmes
 - Define current data system
 - Identify data system gaps and mitigation plans
 - Track artifacts in CM tools

Deploy and scale multiple tiers of MIIC software onto ...



Real hardware at the ASDC and NCDC

(image below from proposal)



Agenda – May 28 Morning

(Building 1250, rm. 161)

Topic	Speaker	Start	End	Duration
Objectives and Requirements	C. Currey	9:00	10:00	1 hour
ASDC Data System	P. Rinsland / M.Little	10:00	10:30	30 min.
NCDC Data System	A. Hall/ J. Morris	10:30	11:00	30 min.
MIIC Software	A. Bartle	11:00	11:30	30 min.
Collaborative Tools	A. Bartle	11:30	12:00	30 min.

Agenda – May 28 Afternoon

(Building 1250, rm. 161)

Working Group Sessions	Start	End	Duration
ASDC MIIC Integration	1:00	2:00	1 hour
NCDC MIIC Integration	2:00	3:00	1 hour
Space Act Draft Review	3:00	3:30	30 min.
Summary, Action Items	3:30	4:00	30 min.

How to intercompare observations

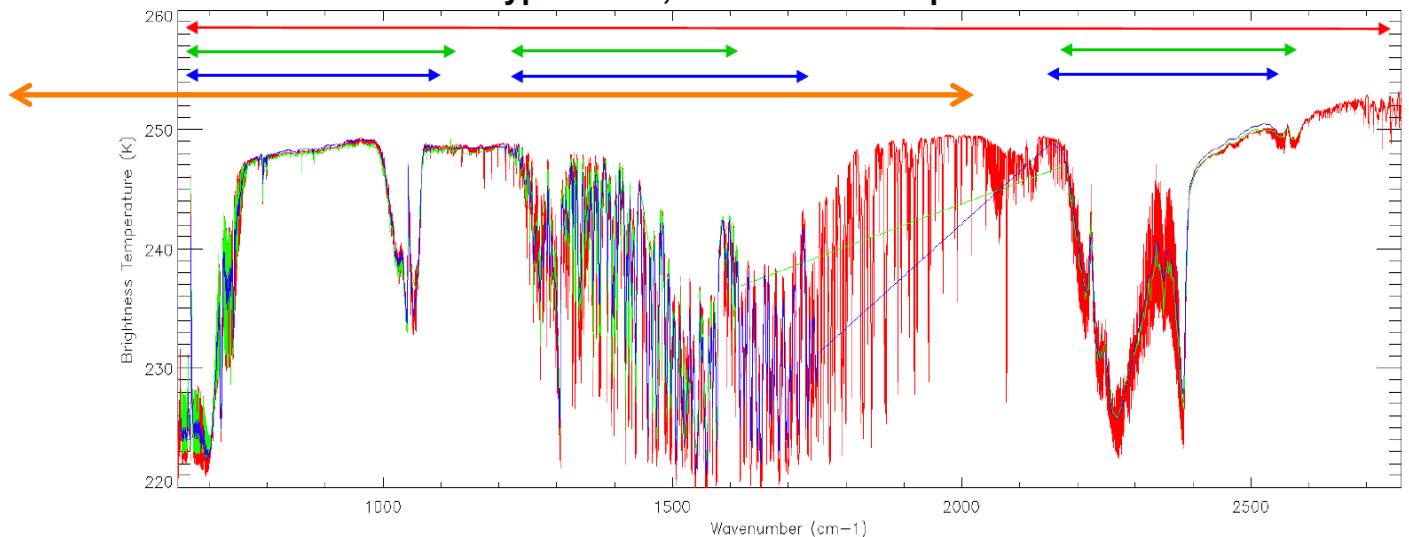
- Match observations in time ($\Delta t \sim 20$ min.)
- Match observations in space (collocation)
 - Grid (0.5°)
 - Nearest neighbor
 - Spatial convolution onto larger FOV
 - Boundary collection of native resolution footprints
- Match observations in spectral range
 - Spectral convolution
 - Spectral resampling
- Match observations in view angles
 - Viewing zenith
 - Solar zenith
 - Relative azimuth

Hyperspectral references for narrowband and broadband instruments

CLARREO LW
200- 2000 cm⁻¹

Instrument	IASI-A	IASI-B	AIRS	CRIS
Satellite	Metop-A	Metop-B	Aqua	NPP
Launch date	2006	2012	2002	2011
Local time	21h30		13h30	
Techno	FTS	Grating	FTS	
Spatial resolution (nadir)	12 km		14 km	
Spectral range	645 – 2760 cm ⁻¹ / 3.62 – 15.5 μm			
Number of channels	8461	2378	1305	
Spectral coverage	Continuous		Partial	
Spectral resolution	0.5 cm ⁻¹	0.4 – 2.1 cm ⁻¹	0.625 – 2.5 cm ⁻¹	

Typical IASI, AIRS and CRIS spectra

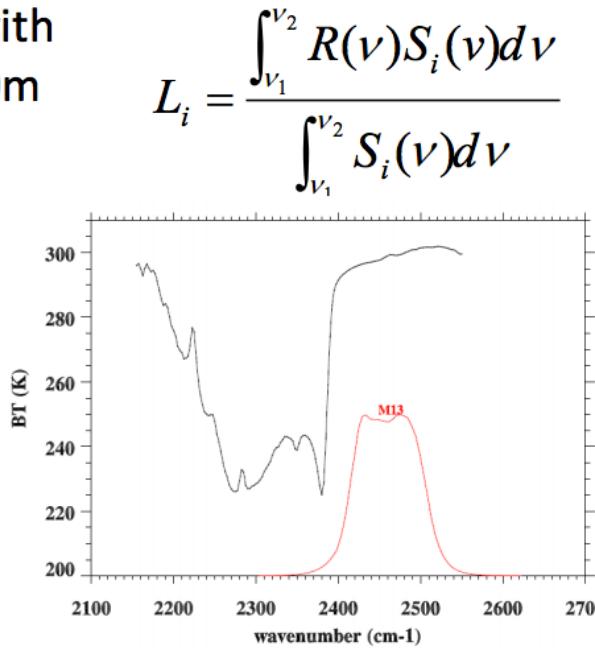
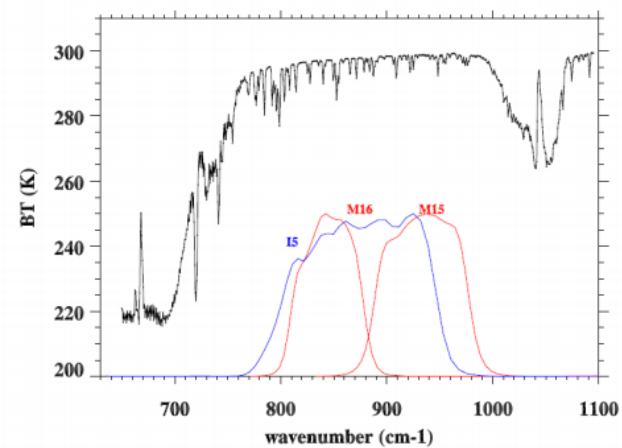


- Hyperspectral reference instruments for LW imager channels
- Image from *D. Jouget (2013), CNES, Radiometric inter-comparison of IASI*

Spectral Convolution to match channels

(CrIS and VIIRS example)

CrIS spectrum is convolved with
VIIRS SRFs for I5 band (350m
spatial resolution)



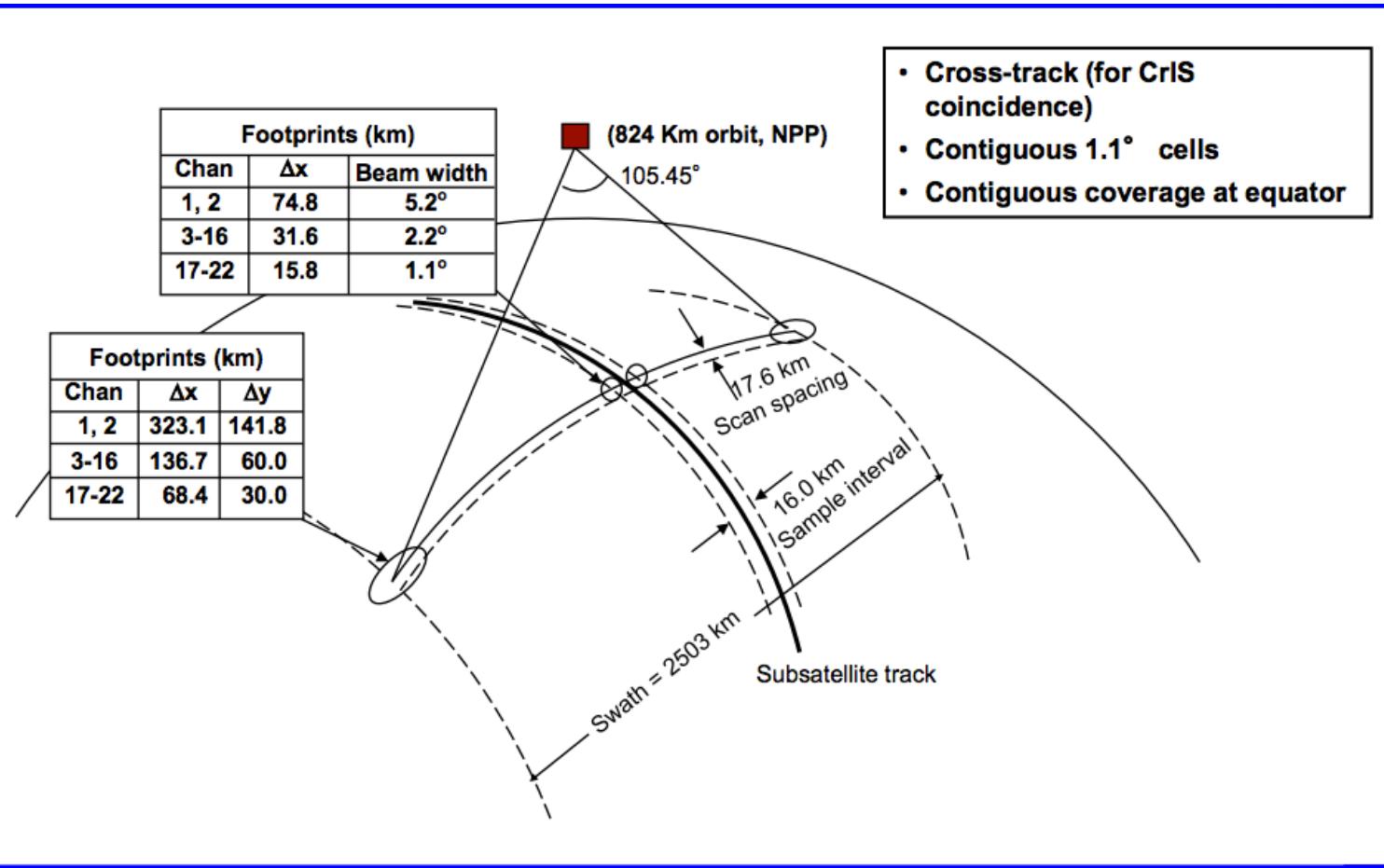
Band No.	Driving EDR(s)	Spectral Range (um)	Horiz Sample Interval (km) (track x Scan)	
			Nadir	End of Scan
M1	Ocean Color Aerosol	0.402 - 0.422	0.742 x 0.259	1.60 x 1.58
M2	Ocean Color Aerosol	0.436 - 0.454	0.742 x 0.259	1.60 x 1.58
M3	Ocean Color Aerosol	0.478 - 0.498	0.742 x 0.259	1.60 x 1.58
M4	Ocean Color Aerosol	0.545 - 0.565	0.742 x 0.259	1.60 x 1.58
I1	Imagery EDR	0.600 - 0.680	0.371 x 0.387	0.80 x 0.789
M6	Ocean Color Aerosol	0.662 - 0.682	0.742 x 0.259	1.60 x 1.58
M6	Atmosph. Correct.	0.739 - 0.754	0.742 x 0.776	1.60 x 1.58
I2	NDVI	0.846 - 0.885	0.371 x 0.387	0.80 x 0.789
M7	Ocean Color Aerosol	0.846 - 0.885	0.742 x 0.259	1.60 x 1.58
M8	Cloud Particle Size	1.230 - 1.250	0.742 x 0.776	1.60 x 1.58
M9	Cirrus/Cloud Cover	1.371 - 1.388	0.742 x 0.776	1.60 x 1.58
I3	Binary Snow Map	1.580 - 1.640	0.371 x 0.387	0.80 x 0.789
M10	Snow Fraction	1.580 - 1.640	0.742 x 0.776	1.60 x 1.58
M11	Clouds	2.225 - 2.275	0.742 x 0.776	1.60 x 1.58
I4	Imagery Clouds	3.550 - 3.930	0.371 x 0.387	0.80 x 0.789
M12	SST	3.860 - 3.840	0.742 x 0.776	1.60 x 1.58
M13	SST Fires	3.973 - 4.128	0.742 x 0.259	1.60 x 1.58
M14	Cloud Top Properties	8.400 - 8.700	0.742 x 0.776	1.60 x 1.58
M15	SST	10.263 - 11.263	0.742 x 0.776	1.60 x 1.58
I5	Cloud Imagery	10.500 - 12.400	0.371 x 0.387	0.80 x 0.789
M16	SST	11.538 - 12.488	0.742 x 0.776	1.60 x 1.58

Table 1: VIIRS Channels
(courtesy of H. Oudrari)

- Generalize Build 1 server-side function (SCIAMACHY) to accept multiple SRFs
- Implement spectral convolution as 2DHisto and N-Tuple “inner” functions
- Will work on any hyperspectral dataset we serve (**CrIS** and **CLARREO**)
- Image from *L. Wang (2013), Inter-Comparison of NPP CrIS with Metop-A IASI*

ATMS FOV Scan Pattern

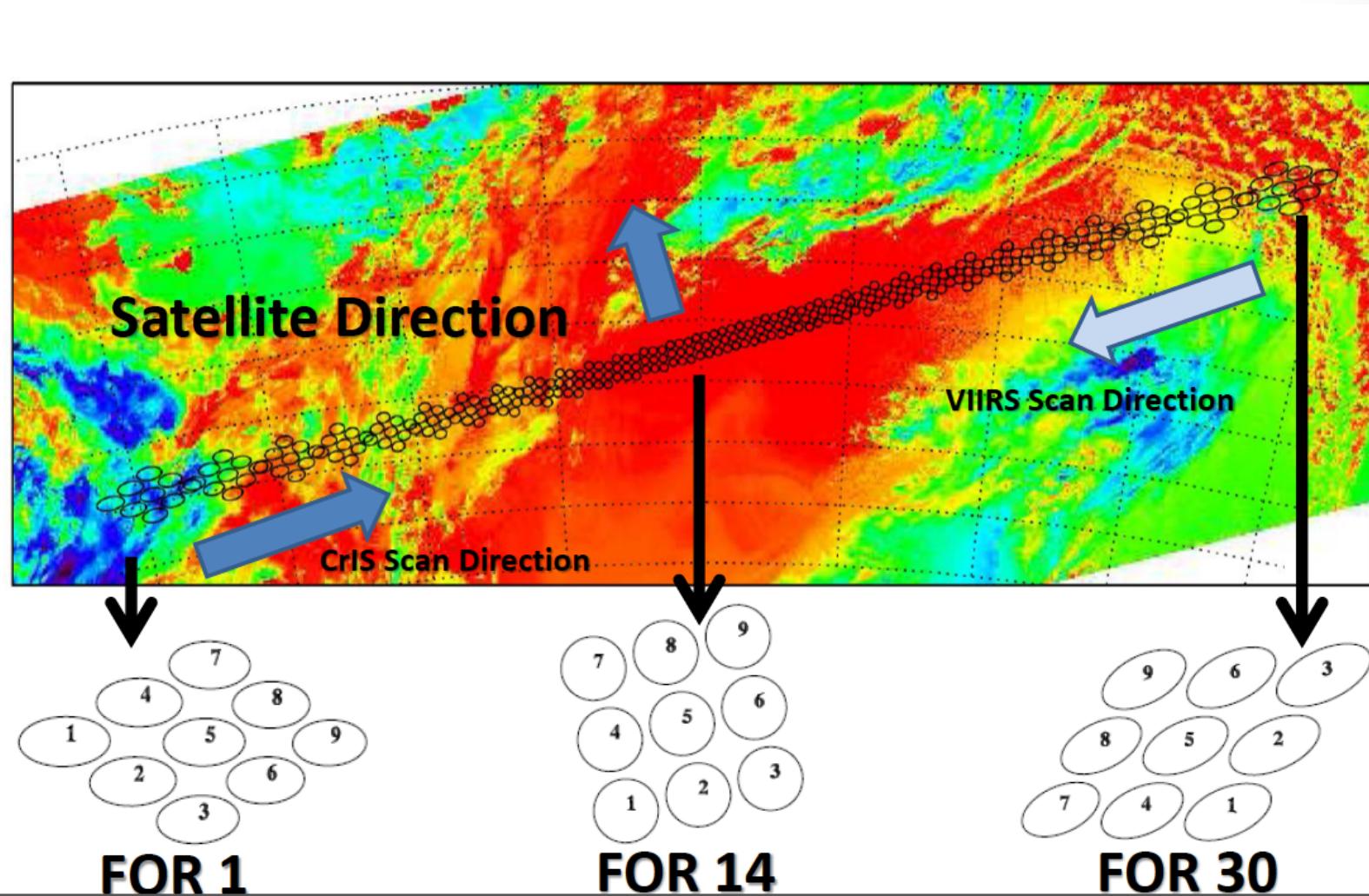
(typical of all scanners – footprint sizes grow with scan angle)



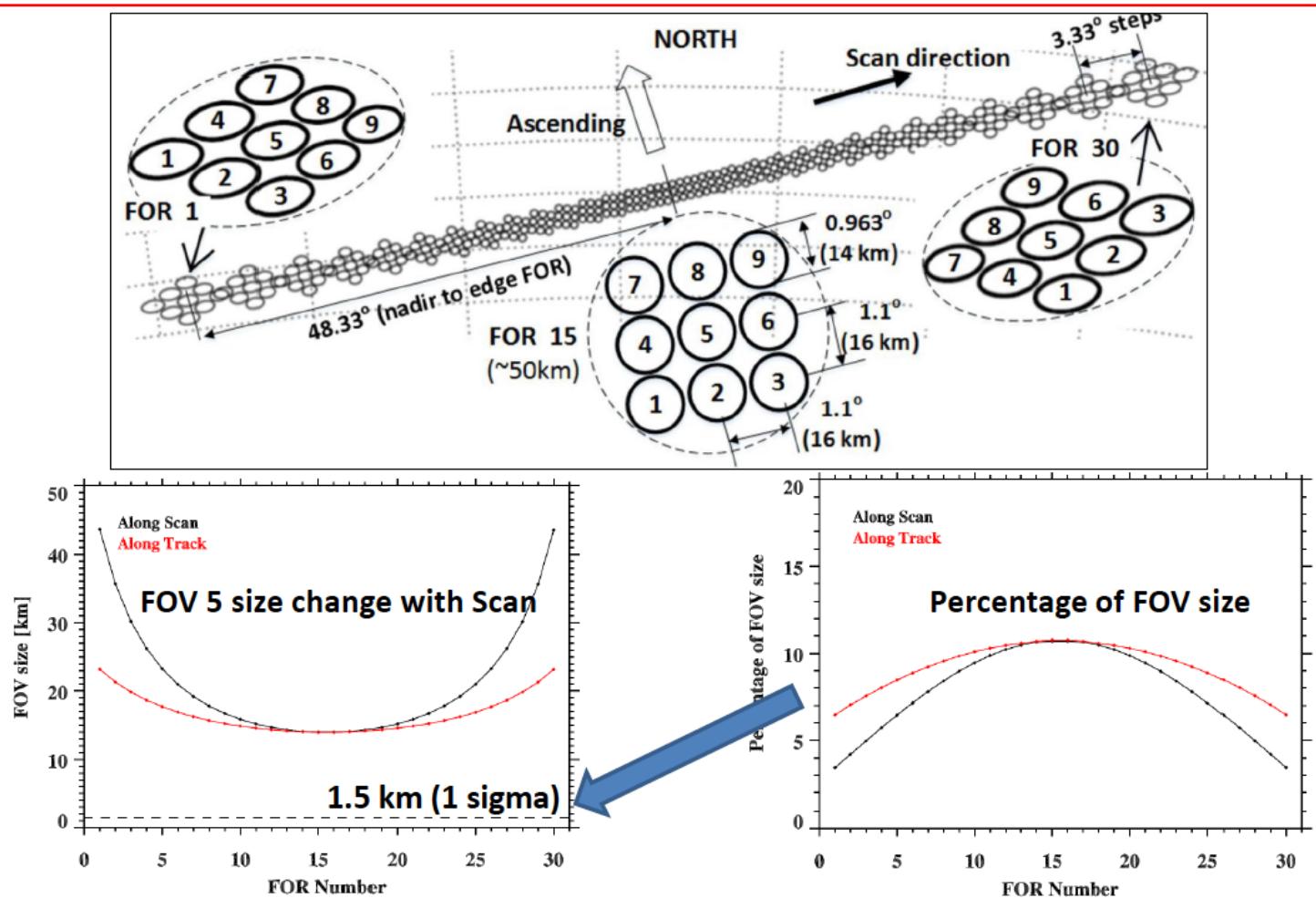
ATMS scan pattern showing how instrument FOV size grows from nadir to 52.7 degrees; image from *R. Leslie (2010), Development and Predicted Performance of the ATMS for NPOESS*

Spatial Convolution

Used to merge imager pixels with larger FOVs



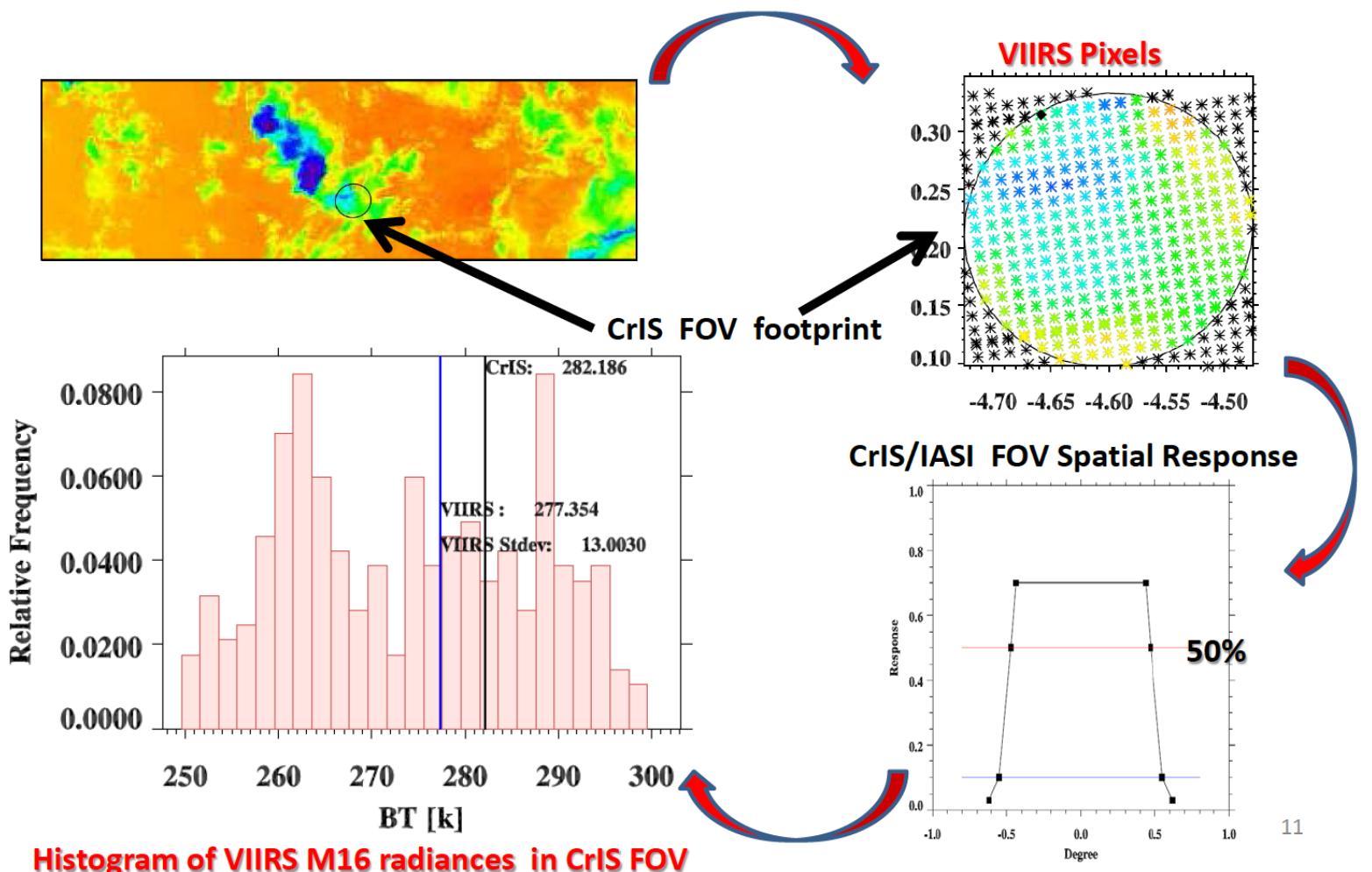
CrIS FOV Scan Pattern



L. Wang (2013) *Suomi NPP CrIS On-orbit Geometric Calibration Performance*

Spatial Convolution

(Example VIIRS with CrIS)



- L. Wang (2013) *Suomi NPP CrIS On-orbit Geometric Calibration Performance*;
- High resolution pixels good for selecting uniform scenes

Collocation Techniques

Need to select/schedule implementations

- #1 Spatially Convolve large FOVs with smaller pixels
 - Use geometric shapes (circle, ellipse) in Earth coordinates that vary with scan angle (Tobin 2006, Nagle and Holz 2009)
 - Use PSF in instrument coordinates (Green 1997, Schreier 2010)
 - Use quasi-conical approach (Nagle and Holz 2009)
 - Examples: CrIS vs. VIIRS, CERES vs. VIIRS, SCIAMACHY vs. MODIS, ...
- #2 Match closest FOVs of similar size
 - *Nearest neighbor* – pair FOVs by distance
 - Examples: CrIS vs. CLARREO, ATMS vs. CrIS, ...
- #3 Group FOVs within spatial boundary
 - Return footprints of varying sizes at native resolution, N-Tuple

MIIC-2 Framework Objectives

- Access matched observations from SNOs and geographic boundaries within large datasets distributed across multi-agency international data centers
- Limit network data transfer from data centers
- Demonstrate successful deployment of a NOAA NASA collaborative distributed data system
- Provide tool to support GSICS and instrument Cal-Val teams
- Enable researchers to more easily access data for studies requiring matched data from multiple instruments
- Demonstrate benefit of using OPeNDAP networking middleware with server-side functions ✓
- Demonstrate feasibility of supporting climate model and empirical data comparisons

MIIC-2 Functional Requirements

- Intercalibration
 - Acquire and analyze matched L1 data at orbit crossings (SNOs)
 - Example: LEO-GEO, NPP VIIRS vs. GOES13. 0.5° grid
- Intercomparison
 - Acquire and analyze matched L1 data at orbit crossings (SNOs)
 - Example: NPP CriMSS EDR (T, H_2O) Cal/Val (*Xu Liu*)
 - CrIS SDR, ATMS SDR, VIIRS EDR, CriMSS EDR, L2 CALIOP cloud products
 - N-Tuple, native res., offline RTM and cloud properties validation
- Data Mining
 - Acquire and filter L1 and L2 data over geographic surface sites (eg., ARM, Libya-4, Dome C,)
 - Mine all supported data products at native resolution, N-tuple (*C. Lukashin*)
 - Surface site event prediction algorithm (*C. Roithmayr*)
- OSSE/GCM Data Mining
 - **Demonstrate** access of large hyperspectral multi-year L3 OSSE datasets with spectral resampling and spatial subsetting server-side functions
 - OSSE vs. SCIAMACHY analysis done offline (*Y. Roberts, D. Feldman*)

All our analyses are types of intercomparison; use generic classes to support all functions

Type Analysis	Target (FOV size)	Reference (FOV size)	Data Products	Parameters	Server-side	Collocation
LEO-LEO or LEO-GEO (SNO)	All Instruments (1 km – 50 km)	All Instruments (1 km – 50 km)	L1 & L2 CERES SSF, VIIRS, CrIS, MODIS, CLARREO, GOES-13, CALIPSO, ATMS, CriMSS	Spectra, Clouds, Flux, Radiance	N-Tuple and 2DHisto	None, native FOVS; Lat-Lon grid (0.5°)
Data Mining (surface site)	All Instruments (1 km – 50 km)	All Instruments (1 km – 50 km)	L1 & L2 CERES SSF, VIIRS, CrIS, MODIS, CLARREO, GOES-13, CALIPSO, ATMS, CriMSS	Spectra, Clouds, Flux, Radiance	N-Tuple and 2DHisto	None, native FOVS; Lat-Lon grid (0.5°)
LEO-LEO (SNO)	Aqua CERES (20 km)	NPP VIIRS (1 km)	CERES L2 SSF, VIIRS L2 Clouds	Clouds	2DHisto	Lat-Lon grid (0.5°)
LEO-LEO (SNO)	Aqua CERES (20 km)	NPP VIIRS (1 km)	CERES L2 SSF, VIIRS L2 Clouds	Clouds	N-Tuple – CERES, SpatConv - VIIRS	Spatial Convolution PSF
LEO-LEO (SNO)	NPP CrIS (14 km)	Aqua MODIS (1 km)	L1 CrIS, L1 MODIS	Apply MODIS RSRs (bands 20-36) to CrIS	SpecConv – CrIS*, SpatConv - MODIS	Spatial Convolution Ellipse
LEO-LEO (SNO)	NPP CrIS (14 km)	Aqua MODIS (1 km)	L1 CrIS, L1 MODIS	Apply MODIS RSRs (bands 20-36) to CrIS	SpecConv – CrIS*, 2DHisto - MODIS	Lat-Lon grid (0.5°)
LEO-LEO (SNO)	NPP CrIS (14 km)	CLARREO (50 km)	L1 CrIS, L1 CLARREO	Brightness Temp.	SpecResamp – CLARREO*, N-Tuple - CrIS	Nearest neighbor match in App Tier
LEO-GEO (SNO)	GOES-13 (4 km)	NPP VIIRS (1 km)	L1 GOES-13, L1 VIIRS	Radiance	2DHisto	Lat-Lon grid (0.5°)
LEO-GEO (SNO)	GOES-13 (4 km)	NPP VIIRS (1 km)	L1 GOES-13, L1 VIIRS	Radiance	N-Tuple	None, native FOVS
LEO-LEO (SNO)	NPP VIIRS, CrIS, CriMSS, ATMS (1 km – 50 km)	CALIPSO (1 km – 5 km)	L2 CALIPSO Clouds, L1 ATMS, L1 CrIS, L2 VIIRS, L2 CriMSS	T, H20, Clouds	N-Tuple	None, native FOVS
OSSE Access (grid boxes)	OSSE	SCIAMACHY (offline)	L3 OSSE	Spectra	SpecResamp – OSSE**	None, gridded

Data Product and Server-side Function Requirements

How do we match observations?

- Step 1 Event Prediction (EP) to define boundaries
- Step 2 Server-side filtering using EP constraints, range checking, and optional spatial and spectral convolution
- Step 3 Application Tier filtering on data returned from server-side functions

MIIC-1 Event Prediction

(C. Roithmayr)

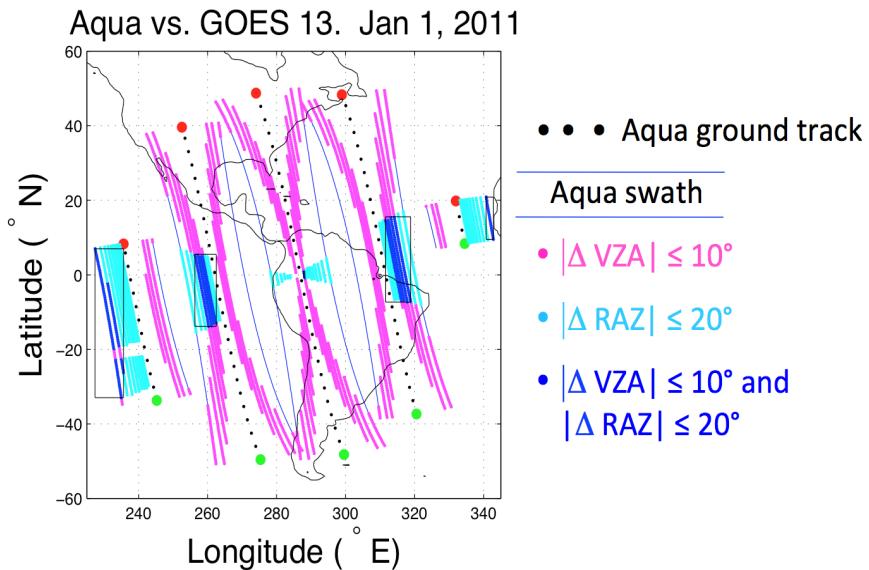


Figure 1. LEO-GEO event prediction showing multiple orbits of the LEO instrument (Aqua MODIS) where observations have matching viewing zenith and relative azimuth angles with the GOES-13 instrument.

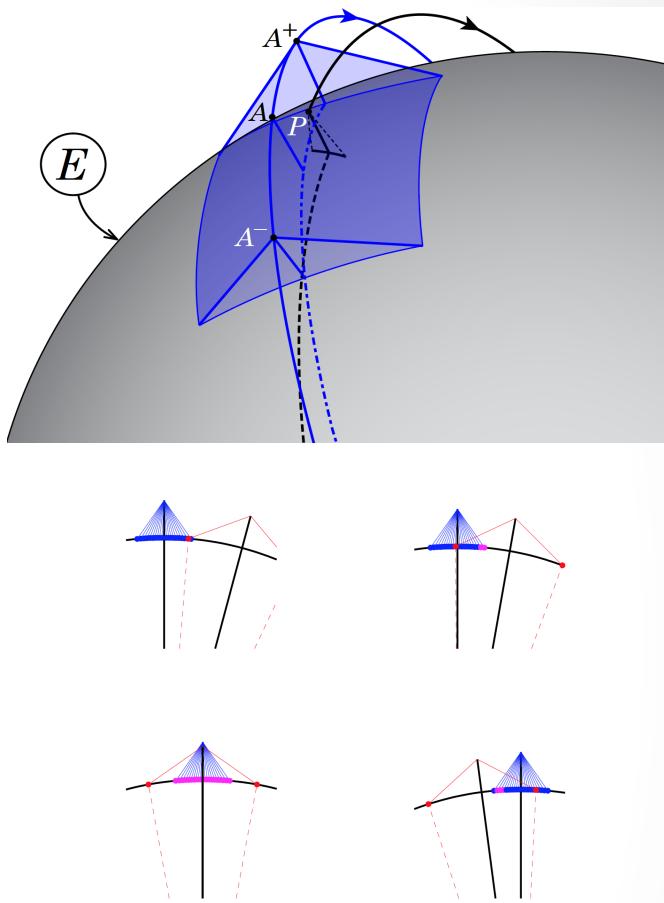


Figure 2. LEO-LEO event prediction creates a tent structure that moves with one spacecraft and predicts the viewing conditions when the second instrument on the other spacecraft is inside the tent (top figure). The bottom figure shows the FOVs with matching view zenith angles (purple) from the two scanners. Demonstrated with Envisat SCIAMACHY and Aqua MODIS data.

MIIC-2 to expand event prediction to mine surface sites (Lacherade et al., 2013), 1° x 1° sites, relax time simultaneity

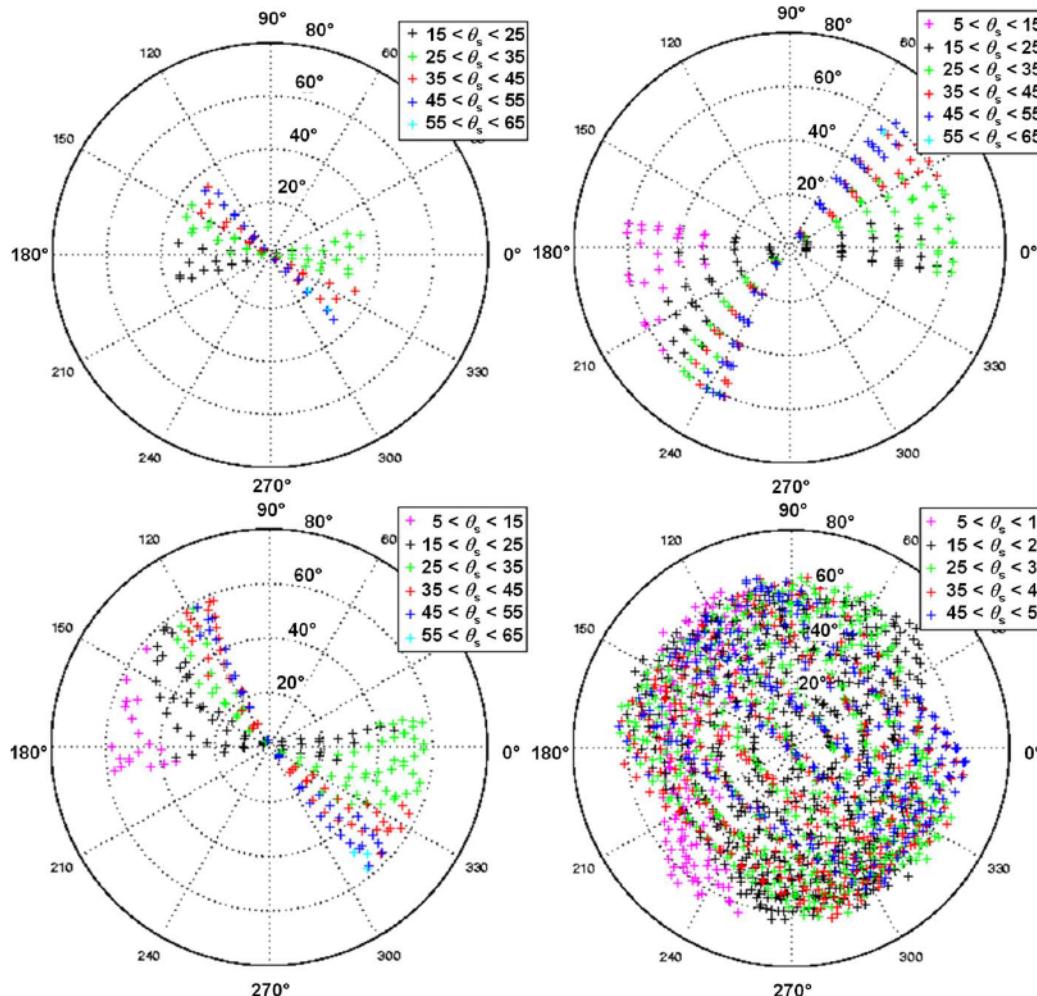
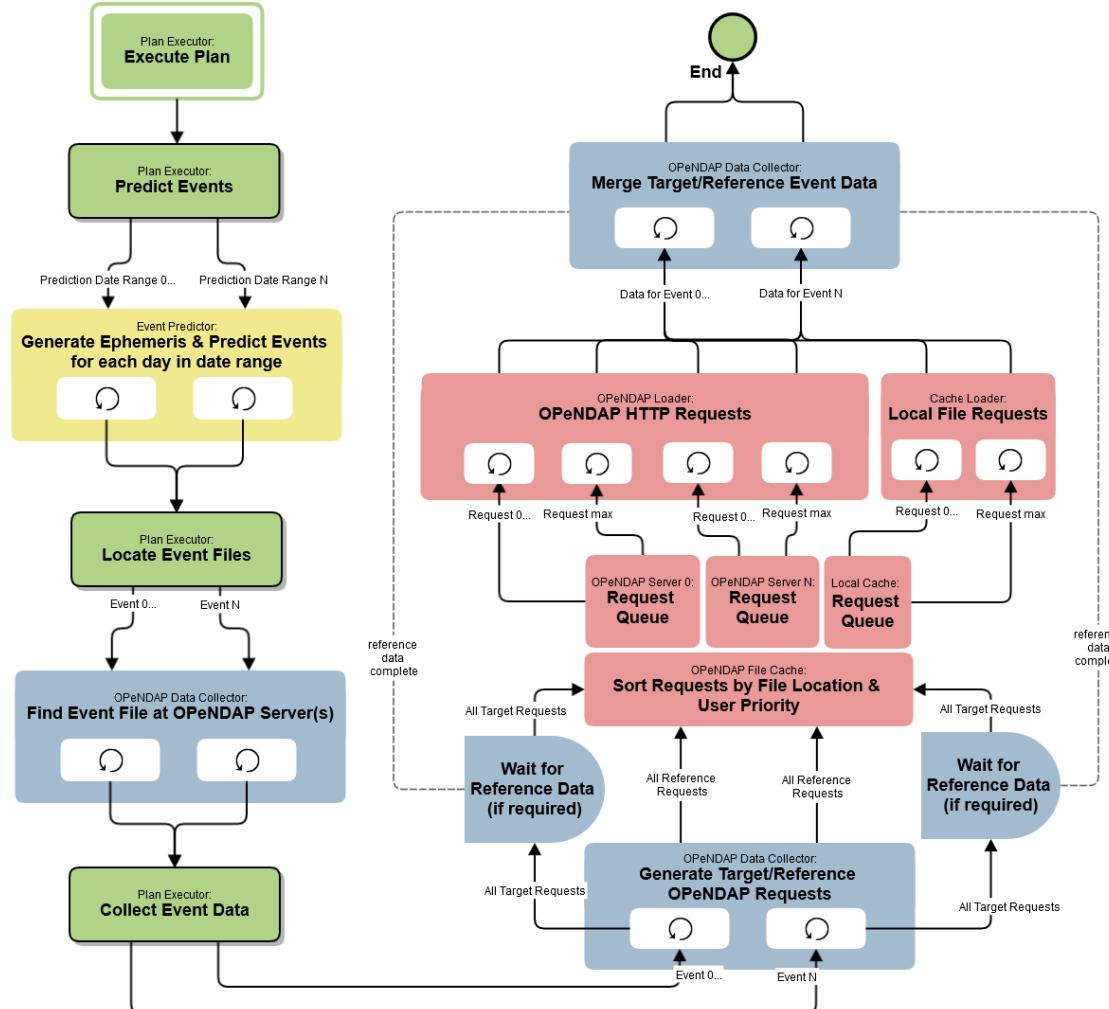


Fig. 7. Typical geometrical sampling of selected measurements for 1 year of acquisition over Libya 1 and for various sensors: MERIS/ENVISAT (upper left), MODIS/AQUA (upper right), VEGETATION/SPOT 5 (lower left), and POLDER/PARASOL (lower right). A color scale is used to classify acquisitions by solar zenith angle ranges of 10°, from 5°–15° (pink) to 55°–65° (cyan). The viewing zenith angle varies from 0° to 80° (nadir corresponds to the center of the polar plot, and 80° corresponds to the external circle), whereas the relative azimuth angle varies counterclockwise, from 0° (on the right) to 360°.

MIIC acquisition software in place, but ... need to stabilize, scale, optimize, and extend functions



The *ICPlanExecutor* sequences the prediction, data acquisition, and analysis services for each *ICPlan*. Where feasible parallelism is used.

MIIC Analysis Services

- The analysis web services to expand significantly!
- Open source histogram analysis and plotting (SCAVIS or JAIDA)
- Incorporate in both REST API and web page UI

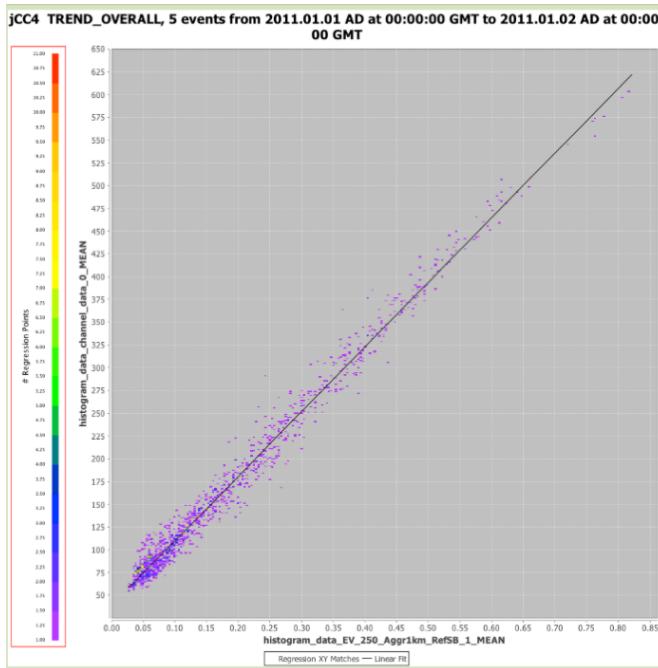


Figure 1. **LEO-GEO**, Jan. 1, 2011 intercalibration of Aqua MODIS band 1 with GOES-13 using 0.5° equal angle grids, 5 overpasses, $vzmax = 10$, $razmax = 20$, $dt = 10$, $R^2 = 0.983$, gain = 709.85, offset = 38.84, N = 1320.

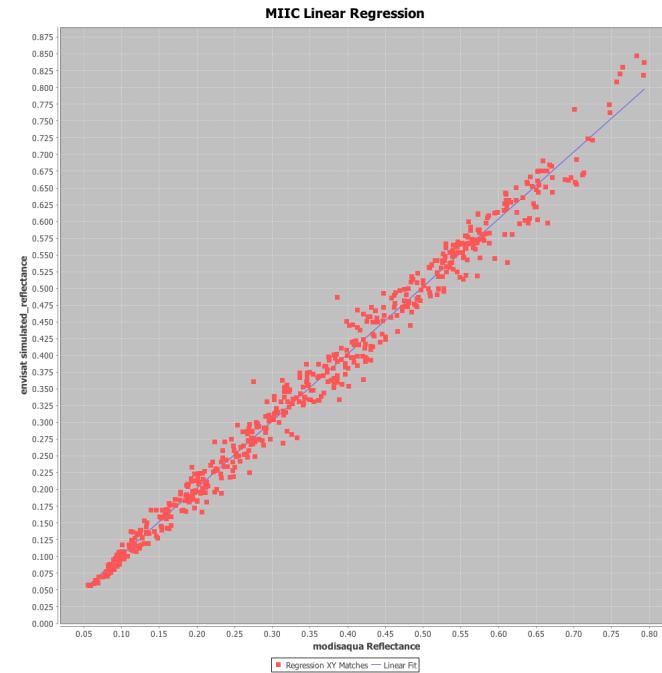


Figure 2. **LEO-LEO**, July 2010 intercalibration of Aqua MODIS band 1 with Envisat SCIAMACHY hyperspectral using spectral and spatial convolution server-side functions. App tier viewing zenith filter (2.0° – 90.0°), $R^2 = 0.987$, gain = 1.004, offset = 0.0013, N = 576.

Algorithm Issues

- Generalize 2DHisto and N-Tuple to work with any data product
- Implement spectral convolution and resampling as 2DHisto and N-Tuple “inner” functions
 - 2DHisto(axisDef(...), axisDef(...), var, *spectralConv*(...), filter, filter)
 - NTuple(var, var, *spectralResample*(...), filter, filter)
- Implement Spatial Collocation as “outer” function
 - *SpatialConvolve*(spatialDef(...), var, var, filter)
 - Define algorithm details – CERES PSF + Ellipse + Nearest Neighbor
- Revisit static (IGBP) and dynamic (% cloudy) mask filtering

Algorithm Issues (cont.)

Dynamic and Static Filter Masks (section 1.3.4.C)

- Proposal: “For each target region defined by the Event Predictor the framework will provide an option to generate a **filter mask** based on dynamic and static data contained within the **L2 CERES** Single Satellite Footprint TOA/Surface Fluxes and Clouds (SSF) data product. The CERES scene type parameter (12 types, integer flag) will be used to build the mask. The scene types are built from static ancillary surface maps (eg., IGBP) and dynamic cloud properties retrieved from CERES measurements. This effort will require collaboration with OPeNDAP to modify the Hyrax server software to allow **HTTP POST** commands to pass in mask data to user provided server-side functions.”
- Proposal: “**demonstrate IGBP filtering on one year of VIIRS/GOES-13 LEO-GEO intercalibration**” – IC based on 0.5° grid
- Not sure this “feature extension” as stated makes sense. To accurately filter VIIRS radiance values could use the L2 VIIRS cloud product and other VIIRS product that contains scene type (TBC). Use VIIRS scene type statistics to filter matched grid boxes. Another approach is to spatially convolve NPP VIIRS and GOES-13 with NPP L2 CERES at NPP/GOES SNOs.
- Revisit requirement to pass mask into server-side function

Algorithm Requirements (cont.)

- Scrub proposal for requirements to be implemented
- Ingest MIIC-2 requirements into CM tools (Confluence?)
- Requirements must be well defined so that software tasks can be defined and tracked (Jira)

MIIC-2 Deployment Issues

- App tier cache(s) – size, shared vs. dedicated cache per node, cache time?
- How best to return data to user? Permanently save any matched data or analysis results?
- Need to optimize database performance – scrub contents of database, only persist what is needed in the future, eg., to collect ESDIS metrics?
- Need dedicated database node that all MIIC servlets can access?
- Configure proxy servers to support multiple MIIC servlets (Tomcat)?
- How to route HTTP requests from App Tier to NCDC OPeNDAP server?
- Move/refactor TLE Server software into App Tier (eliminate delay, reliability); run on which node?
- Evaluate current data system performance and specify goals
 - Number of concurrent users, response times per web service request type, ...
 - Monitor OPeNDAP and App tier performance, identify/fix bottlenecks
- Define how to track issues for all major components in CM tools
-

Components to track

- UI tier - REST API and web page
- App tier
- OPeNDAP tier and server-side functions
- TLE server
- Local data tier – database and cache
- ASDC hardware and MIIC software integration
- NCDC hardware and MIIC software integration

Need to compare current performance to Build 1 LEO-GEO

Test	OPeNDAP Servers (Data)	Staged Input #Files (GB)*	Transfer DODS Files (GB)	Reduction Factors	Bes Server Resident Memory (MB)	Download Time secs. (mins.)
OPeNDAP 1 server	dataserver2 (G-13, MODIS L1B)	9672 (1357)	808** (62.1)	12.0 (21.9)	161 MB, 141 MB	2216 (36.9)
EA Grid 1 server	dataserver2 (G-13, MODIS L1B)	9672 (1357)	808 (1.8)	12.0 (753.9)	143 MB, 143 MB	3974 (66.2)

Table 1. Build 1 performance – one month MODIS L1B vs. GOES-13, original equal angle SSF (prior to 2DHisto)

*MODIS L1B files (storage) per month: 8928 files (1339 GB); GEO 744 files (18 GB)

**130 GEO (6.1 GB) 678 MODIS (56 GB)

Client filtering: -vza 15 -raz 15

1-3 parameters gridded each instrument

- Processed one month, January 2011
- 98% time spent in server-side functions + network transfer, 2% in client processing
- Prediction accounts for 22x reduction in data volume transferred
- Gridding accounts for an additional 34x reduction
- Build 1 total reduction factor of 754

Performance

- Aron to present current performance
- 2DHisto out of memory (OOM) error when select all 97 science parameters within L1B MODIS
- Major refactor required to solve problem